PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/16, 15/62, C12P 21/02 C12N 1/21, 5/16 // (C12N 1/21 C12R 1/19)

A1

(11) International Publication Number: WO 92/06194

(43) International Publication Date: 16 April 1992 (16.04.92)

(21) International Application Number: PCT/IT91/00079

(22) International Filing Date: 26 September 1991 (26.09.91)

(30) Priority data:
48315 A/90 27 September 1990 (27.09.90) IT

(71) Applicant (*for all designated States except US*): CONSIGLIO NAZIONALE DELLE RICERCHE [IT/IT]; P.le Aldo Moro, 7, I-00185 Rome (IT).

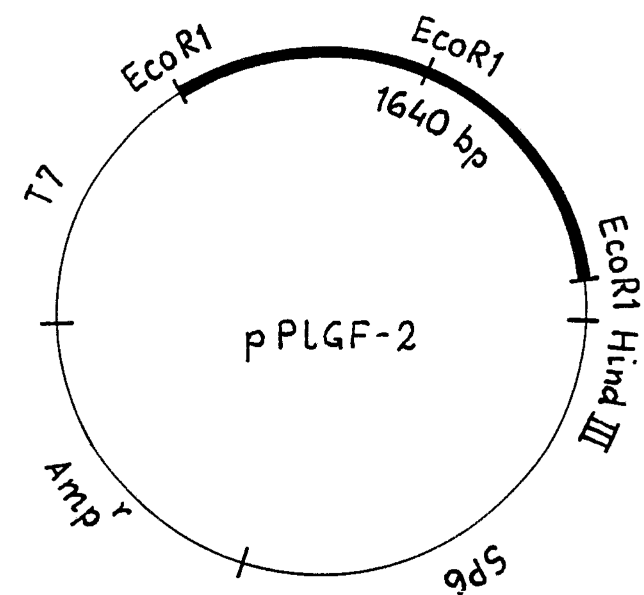
(72) Inventors; and
(75) Inventors/Applicants (*for US only*) : PERSICO, Maria [IT/IT]; MAGLIONE, Domenico [IT/IT]; I.I.G.B. C.N.R., V.le Marconi, 10, I-80125 Napoli (IT).

(74) Agents: BANCHETTI, Marina et al.; Ing. Barzanò & Zanardo Roma S.p.A., Via Piemonte, 26, I-00187 Rome (IT).

(81) Designated States: AT, AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH, CH (European patent), CI (OAPI patent), CM (OAPI patent), CS, DE, DE (European patent), DK, DK (European patent), ES, ES (European patent), FI, FR (European patent), GA (OAPI patent), GB, GB (European patent), GN (OAPI patent), GR (European patent), HU, IT (European patent), JP, KP, KR, LK, LU, LU (European patent), MC, MG, ML (OAPI patent), MN, MR (OAPI patent), MW, NL, NL (European patent), NO, PL, RO, SD, SE, SE (European patent), SN (OAPI patent), SU+,TD (OAPI patent), TG (OAPI patent), US.

Published
With international search report.

(54) Title: NUCLEOTIDE SEQUENCES CODING FOR A HUMAN PROTEIN WITH ANGIOGENESIS REGULATIVE PROPERTIES



(57) Abstract

A cDNA sequence coding for a human protein having regulative properties of angiogenesis has been isolated and sequenced. Expression vector containing such sequence have been constructed and, after transformation of host cells, the synthesis of the related protein has been obtained. Such protein called PIGF can be employed both in the tumoral immunologic and diagnostic field and in the therapeutic field for pathologies related to the formation of vessels, as for instance the healing of wounds and so on.

+ DESIGNATIONS OF "SU"

Any designation of "SU" has effect in the Russian Federation. It is not yet known whether any such designation has effect in other States of the former Soviet Union.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU+	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		

NUCLEOTIDE SEQUENCES CODING FOR A HUMAN PROTEIN WITH ANGIOGENESIS REGULATIVE PROPERTIES

This invention relates to nucleotide sequences coding for a human protein having angiogenesis regulative properties.

More particularly, this invention relates to the isolation and to the molecular characterization of a gene coding for a new protein having the properties of an angiogenic factor which regulates in vivo the formation and/or the regeneration of the vertebrate blood vessel system, and it also relates to the protein itself. Moreover, this invention also refers to vectors containing such sequence or parts thereof, to prokaryotic and eukaryotic cells transformed with such vectors, and to the employment of such vectors and of such cells for the production of the protein and of corresponding polyclonal and/or monoclonal antibodies as well.

It is well known that growth factors are polypeptides, synthesized and secreted by mammalian cells, capable of acting not only on the proliferation, but also on the differentiation and morphogenesis of target cells. Indeed, it has been shown that some growth factors exert their action by regulating mechanisms such as chemiotaxis, activation of inflammatory system cells and repairing of tissues (Whitman, M. and Melton, D. A., 1989, Annual Rev. Cell Biol., 5, 93-117).

Because of the similar phenotype between cultured growth factors stimulated and retrovirus transformed cells, it has been suggested that common mechanisms

control such phenomena. Indeed, the interaction between a growth factor and its own specific receptor indirectly activates gene activity regulative proteins, through intermediate reactions involving different protein-kinases. Many of the components of this metabolic chain have been identified as the cellular analogs of viral oncogenes, suggesting how oncoviruses could interfere with normal cellular processes.

Many growth factors have been identified up to the present time, the corresponding genes have been cloned, and such factors have been divided into groups, on the basis of similar activities and/or of sequence homologies; among them there is the family of angiogenic factors.

Angiogenesis, or the formation of vessels of the vascular system, is a complex process occuring during embryogenesis, wound healing and organ regeneration. Moreover, some pathologies like the growth of solid tumors, some retinopathies and rheumatoid arthritis induce an aberrant angiogenesis (Risau W., 1990, Progress in Growth Factor Research, 2, 71-79).

Angiogenesis _in vivo_ is a multi-step process, two of them being represented by the migration and the proliferation of endothelial cells devoted to the formation of vessels.

In the most recent years, many angiogenic factors have been identified, and the corresponding genes cloned. Among them: angiogenin, subject-matter of the patent application PCT no. 8701372; the platelet-derived endothelial growth factor PD-ECGF (Ishigawa et al., 1989,

Nature, 338, 557); the human vascular permeability factor, VPF (Keck et al., 1989, Science 246, 1309), which was cloned also in the mouse with the denomination of vascular endothelial growth factor, VEGF (Leung et al., 1989, Science, 246, 1306); the growth factors for fibroblasts, i.e., the acid factor, a-FGF, and the basis factor, b-FGF, the transforming growth factors alpha, TGF-.., and beta, TGF-..(Folkman and Klagsburn, 1987, Science, 235, 442).

Angiogenic factors have been divided into two groups, according to their way of action: either directly on the vascular endothelial cells, by stimulating motility or mitosis, or indirectly on cells producing growth factors acting on endothelial cells.

In vitro analysis have put into evidence that angiogenic factors exert different effects on the motility and on the proliferation of endothelial cells. Indeed, some of them stimulate just one of the two events, other ones stimulate both events, whereas others seem to be ineffective *in vitro*, and ,lastly, other ones show even an inhibiting activity of the endothelial cellular proliferation. Such data point out that the regulation of angiogenesis is a complex process mediated by diffrent components, many of which have not been identified as yet.

Accordingly it is evident that there is the need for identifying and isolating new angiogenic factors capable of stimulating the migration and differentiation of endothelial cells, to be utilized both in the diagnostic field, as tumoral markers and for inflammatory

diseases, and in the therapeutical field, for topic or internal use, for instance in the treatment of wounds, of tissues after a surgical operation, of transplantation, of burns, ulcers, etc.. Such factors can be employed successfully also *in vitro*, as growth-stimulating of cell cultures.

Moreover, DNA recombinant techniques allow such factors to be produced in suitable amounts, in short times and at remarkably low costs.

Indeed, there is an increasing need for identifying new specific tumoral markers because of uncertainties in tumor diagnosis. Moreover, recent methods for producing hybrid proteins (Fitzgerald D. and Pastan I., 1989, J. Natl. Cancer Inst. 81, 1455-1463) and/or conjugate antibodies (Pearson, J. W. et al., 1989 Cancer Res. 49, 3562-3567) with toxic molecules, are giving promising results in the field of tumoral serotherapy, with an increasingly growing demand for new factors to test. Finally, many of angiogenic factor genes have been cloned starting from tumoral cells, whereas a better applicability in the therapeutic field of genes coming from non-neoplastic material is evident.

Accordingly, this invention provides nucleotide sequences coding for a protein having a regulative activity of angiogenesis, said sequences being obtained from non-neoplastic tissue; vectors containing said sequences; cells transformed by said vectors and producing protein having biologic and/or immunologic activities of a new angiogenic factor, as well as the protein itself, to be employed in diagnostic and

therapeutic fields.

This invention also provides a procedure for obtaining the protein, or parts thereof, by recombinant techniques, as well as its use as an antigen for the production of the corresponding polyclonal or monoclonal antibodies.

Indeed, molecular probes comprising sequences coding for the angiogenic factor subject-matter of the present invention can be employed as markers in the diagnosis of pathologies related to the aberrant production thereof, as the case of some tumoral pathologies for other angiogenic factors.

Moreover, the protein which is another subject-matter of this invention can be employed in the treatment of inflammatory diseases, in the treatment of wounds, of tissues after surgical operations, of transplantation, of burns of ulcers and so on. Such factor can also be employed in vitro successfully, as growth stimulating of cell cultures.

Finally, DNA recombinant techniques employed in the present invention allow to produce the molecular probes and proteins described above in suitable amounts, in short times and at remarkably reduced costs.

The nucleotide and amino acids chains of this invention can be employed for diagnostic tests and for theraupetic purposes, both as directly derived from host cells and as after suitable modifications, for obtaining a better production for compositions.

Accordingly, the object of this invention consists in nucleotide sequences coding for a protein,named PlGF,

with immunogenic and/or biologic properties of an angiogenesis regulative factor, having the amino acids sequence of SEQ ID N1.

As another embodiment of the invention, the PlGF amino acid sequence derives from alternative splicings of the primary transcript, preferably at the nucleotide sequence shown in SEQ ID N2, most preferably giving rise to an amino acid insertion of 21 amino acids, whose sequence is shown in SEQ ID N2, at position 141-142 of the amino acid sequence shown in SEQ ID N1.

An object of this invention also consists in nucleotide sequences coding for the PlGF protein, lacking and/or substituted in one or more amino acids, preferably deleted from the amino acid 1 to the amino acid 31 of SEQ ID N1; the present invention also provides nucleotide sequences which are allelic derivatives of the sequence coding for SEQ ID N1, as well as nucleotide sequences that are complementary to those coding for SEQ ID N1.

Again according to the present invention the nucleotide sequence can be covalently bounded to a nucleotide sequence which can be translated into amino acid sequence by employing the same reading frame of the gene coding for PlGF, which preferably does not interfere with the angiogenesis regulative activity of PlGF, and which more preferably codes for a protein portion having toxic activity.

Accordingly, the object of this invention also consists in the nucleotide sequence of SEQ ID N1, even though the same is lacking and/or substituted in one or more nucleotides, coding at its coding part for the

protein PlGF, as in SEQ ID N1.

The present invention also provides nucleotide sequences hybridizing with SEQ ID N1, or parts thereof; nucleotide sequences obtained both through natural and synthetic or semisynthetic methods, by substitution, deletion, insertion and inversion mutations, either concerning single bases or multiple bases, of sequence described in SEQ ID N1, or parts thereof; and nucleotide sequences comprising sequences coding for a protein having immunogenic and/or biologic properties similar to those exhibited by the protein PlGF or parts thereof.

A further aspect of this invention relates to the protein PlGF having the sequence disclosed in SEQ ID N1, or parts thereof, obtained either by means of recombinant DNA techniques or isolated from biologic tissues. Said protein, or parts thereof which are immunologically active, can be employed as antigenes for producing polyclonal and/or monoclonal antibodies.

The present invention also provides cloning and/or expression vectors, both prokaryotic and eukaryotic, comprising the nucleotide sequences subject-matter of the invention, sequences promoting transcription located upstream and, in general, a selective marker. Preferably, sequences promoting transcription in an inducible manner, can also be present and enhancers, polyadenylation signals and so on, as well.

Again an object of the present invention consists in prokaryotic and eukaryotic cells transformed by said vectors to be employed for producing the PlGF protein or parts therof.

Just for illustrative and not for limitative purposes the present invention will be described in the following examples. In what follows reference will be made to the enclosed Figures wherein:

Figure 1 represents the restriction map of the recombinant λ GT11 phage, comprising the sub 32 fragment;

Figure 2 represents a "Northern blot" experiment employing the sub 32 cDNA fragment;

Figure 3 represents the restriction map of the plasmid pPlGF-2;

Figure 4 represents an exemplificative scheme of subcloning of a fragment coding for a portion of the protein PlGF in the expression vector pET3 (Novagen, Madison WI, USA);

Figure 5 represents a polyacrylamide gel electrophoresis of the protein PlGF, said protein being obtained through the recombinant way.

Example 1

Isolation of the cDNA coding for a new angiogenic factor

A first cDNA fragment, named sub 32 , was isolated from a clone of a cDNA library from human placenta, in the λ GT11 vector, according to conventional procedures and employed also in other laboratories (Wataneb et al., J. Biol. Chem. 264, 12611-19, 1989).

Briefly, RNA was extracted through lysis with guanidine thiocyanate and centrifuging on a discontinuous gradient of caesium (Sambrook J., Fritsch E.F., Maniatis T., Molecular Cloning - A Laboratory Manual. Second edition. Vol. 1, 7.19. Cold Spring Harbor Lab. Press). The poly A+ RNA was purified through chromatography on

oligo-dT cellulose (ibid. 7.26). The cDNA synthesis and the cloning of the λ GT11 phage vector (Stratagene, La Jolla California, USA) in the Eco R1 restriction site was carried out following the protocol described ibid. 8.54-8.79. A clone, whose map is shown in Figure 1, was identified because of comprising also a sequence of 2600 nucleotides capable of hybridizing, in 5x SSC at 65C according to the hybridization procedure on filters described ibid. 8.46, with a sequence coding for the cDNA of the glucose-6-phosphate dehydrogenase enzyme (G6PD) (Persico, M. et al., 1986, Nucl. Acid Res., 14, 2511). A fragment of 240 nucleotides was also isolated from this recombinant phage after digestion with Eco R1 and Bam H1, and the fragment was called sub 32. Said fragment, after labelling with 32P by means of the "nick translation" procedure disclosed ibid. 10.6-10.8 was employed for:

a) analyzing RNAs extracted from different tissues or cell lines, by "Northern blot" procedure as described in íbid. 7.37. The results shown in Figure 2 show that the sub 32 fragment detects specific mRNA in the placenta (line 2), in HEPG2 hepatoma cells, ATCC N. HB8065 (line 3), in JEG human choriocarcinoma cells, ATCC N. HTB36 (line 4) and, at lower concentration, in Hela S3 cells, ATCC N. CCL2.2 (line 5), but not in HL60 cells, ATCC N. CCL240 (line 1);

b) screening a cDNA library from JEG human choriocarcinoma JEG, ATCC N. HTB36, according to the procedures described for the cDNA library from human placenta, in the λ GT10 vector (Stratagene, La Jolla California, USA), in the Eco R1 site. Two clones were

isolated, digested with Eco R1 and subcloned in the pUC 18 vector (Stratagene, La Jolla, California, USA) and the sequence determined by Sanger's method (ibid. 13.6-13.10). The sequence revealed the fragmnents overlapped partially one another, but did not comprise the whole sequence coding for the corresponding mRNA. Hence, the isolated fragments were employed for a second screening, employing the same techniques. The library employed was the cDNA library from human placenta, from which the initial sub 32 fragment came from. Then two clones were isolated, their DNA was digested with Eco R1, the resulting inserts were subcloned in the pGEM 1 vector (Promega Corporation, Madison WI., USA) and their sequence was determined by Sanger's method. The two DNA fragments obtained after digestion with Eco R1 were religated together through T4-ligase and cloned in the same pGEM 1 vector in the Eco R1 site, to obtain the whole cDNA sequence corresponding to the mRNA present in the placenta, in a single plasmid, called pPlGF-2 (ATCC Dep. No. 40892), whose map is shown in Figure 3.

In order to confirm that the resulting fragment covers the whole coding sequence, the sequence was compared with the sequence of a genoma fragment obtained after hybridization of the same fragment with a genomic library from human fibroblasts WI38 (No. 944201 Stratagene, La Jolla, California, USA) in the λ Fix vector.

The cDNA sequence was identified according to Sanger's method (ibid. 13.3-13.10) and revealed:

a) a 5' end untranslated region of 321 nucleotides

comprising a sequence capable of forming a stem-loop secondary structure, indicative of a translation regulative signal;

b) a sequence of 447 nucleotides with an open reading frame coding for a protein of 149 amino acids, comprising a hydrophobic sequence of 32 amino acids at the NH2-terminal, indicative of the signal peptide of secreted proteins;

c) a 3' end untranslated region of 877 nucleotides comprising a polyadenylation site.

The amino acid sequence, deduced by the cDNA sequence, was inserted into the European Molecular Biology Laboratory (EMBL) Data Bank, showing no protein with the same sequence. A 50% homology, limited to a 120 amino acid region, was shown with the vascular permeability factor VPF (Keck et al., 1989, Science, 246, 1309), a powerful angiogenic factor, thus suggesting that the new protein PlGF can have by itself an angiogenesis-regulating activity.

Example 2

<u>Screening of a cDNA library from JEG-3 cells with pPLGF and structure of PlGF-gene</u>

A cDNA library, obtained from JEG-3 cell mRNA, was screened with the PlGF probe. Six recombinant phages were isolated. The sequence of two of them revealed they have a lenght of 510 bp, generating a 170 amino acid protein. The sequence resulted to be identical to the cDNA isolated from placenta, but an insertion of 63 bp, generating a 21 amino acid insertion into the protein, at position 141-142. Interestingly, the new sequence

contains 10 basic amino acids (Arg and Lys) over 21.

Example 3

enomic mapping and cloning of the PlGF gene

The gene coding for the protein PlGF was mapped on the chromosome 14 through "Southern blot" analysis, by employing DNA from different hybrid cellular lines, each containing different human chromosomes (not shown).

The structure and part of the nucleotide sequence of the PlGF gene was determined from a human genomic library. The gene is divided into six exons and five intervening sequences generating, through splicing, the transcript coding for the 149 aa. protein. In choriocarcinoma cells (JEG-3) the primary transcript is alternatively spliced at the fifth intron to generate a transcript coding for the 170 aa. (see SEQ ID N2). Another alternative splicing involving the sequence from 174 to 828 of SEQ ID N2 of the fifth intron, gives rise to an higher molecular weigth PlGF protein. In fact two proteins are immunoprecipitated from JEG-3 conditioned medium, with antibodies anti PlGF.

Example 4

Subcloning of PlGF cDNA in a prokaryotic expression vector

A scheme of the subcloning strategy is shown in Figure 4, wherein the pET3 vector was employed (Novagen Madison Wi; USA) containing essentially the T7 phage RNA polymerase promoter, the terminator of the same phage, an origin of the replication (ori) and the resistance to ampicillin (amp).

The cDNA insert to be subcloned was obtained

through PCR amplification (polymerase chain reaction, ibid. 14.6), generating a cDNA coding for the protein lacking the first 31 amino acids. As template, the Eco RI DNA fragment, from nucleotide 1 to nucleotide 940, was employed. As primers for RNA polymerase the following oligonucleotides were employed, synthesized with an "Applied Biosystem 381A" oligo-synthesizer:

- oligonucleotide A complementary to the coding strand from the nucleotide 768 to the nucleotide 787, in which the GGATCC sequence, Bam H1 recognition site, was inserted between nucleotides 775 and 776, having the following sequence:

5'-TCCTCCAAGG<u>GGATCC</u>TGGGTTAC-3'

BamH1

- oligonuceotide B complementary to the non-coding strand from nucleotide 404 to nucleotide 421, in which the CATATG sequence, Nde 1 recognition site, was inserted between the nucleotides 414 and 415, having the following sequence:

3'-GCAAGGG<u>GTATAC</u>TCGTCTGTTCC-5'

Nde1

The nucleotide chain, obtained from PCR, was digested with Nde 1 and Bam H1 and ligated with the prokaryotic pET3 expression vector in the same Nde 1 and Bam H1 sites according to standard protocols. The product was employed for transforming the E.coli HB101 strain which had been made competent with the CaCl2 method. The recombinant plasmid was identified and employed for transforming the E.coli JM109 strain (DE3, Promega Corporation, Madison WI, USA).

Example 5

Synthesis and isolation of the PlGF protein from bacteria

A single colony was inoculated in LB broth containing 100 μg/ml of amplicillin (Sigma, St. Louis MO., USA) and 4 g/l of glucose and then grown at 37C to reach an optical density O.D. of 0.35 at 600 nm. IPTG (Sigma) was added to a 1 mM final concentration and the culture was incubated at 37C for additional 3 hours. The culture was centrifuged and resuspended in 1/10 of the initial volume of a buffer containing 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 (TE). Following further centrifugation, the precipitate was resuspended in 1/60 of the initial volume into a lysis buffer containing TE, 1% SDS, 0.1 M NaCl. Bacteria were divided into aliquots of 500 ul and subjected to lysis by three cycles of freezing and thawing, followed by middle strength sonication.

An example of the resulting electrophoretic pattern is shown in Figure 5, wherein lines 1, 2, 3 and 4 represent electropnoretic patterns of proteins from lysates respectively 0, 1, 2 and 3 hours after IPTG induction. As control, line 5 represents the same strain transformed only with the vector lacking the insert, induced with IPTG for 3 hours. Electrophoresis was carried out according to Laemli, Nature, 227, 680-685, 1970, in a 15% polyacrylamide gel stained according to the method described by Bradley et al., Anal. Biochem. 182, 157-159 (1989).

Example 6

Production of anti PlGF antibodies and immunoprecipitations of PlGF

70 μg of the protein PlGF was employed for immunizing two chickens, as described by Gassmann et al., 1990 Faseb J. 4, 2528-2532. The antibodies so formed were extracted and purified from the yolk through precipitation with polyethylene glycol (PEG) as described by Gassmann et al. (cf. above). The immunoprecipitations were performed by incubating 120-250 μl of cellular lysate, or Cos-1 cell conditioned medium, with 10 or 15 ul of rabbit or chicken antibodies, for 2 hours at room temperature, or 16 hours at 4C. The immunoreactions with chicken antibodies were further treated with 15 ul of rabbit anti-chicken IgG (SIGMA N. C6778), for 1 hour at room temperature.

The immunocomplex was selectioned through protein-Sepharose 4B (Pharmacia) and washed twice with 1.2 ul of PBS with 0.01% Nonidet-P40 and 400 uM of NaCl.

The immunoprecipitates were then resuspended and analyzed on polyacrylamide gel under denaturing and reducing conditions according to standard procedure. If COS cells had been previously transfected with the plasmid pSVL-PlGF, a protein of 25 KDa molecular weight is immunoprecipitated, both from the lysate and from the culture medium.

SPECIMEN SEQUENCE LISTING

SEQ ID N.1

SEQUENCE TYPE: nucleotide with corresponding protein at the coding region

SEQUENCE LENGTH: 1645 bp

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHETICAL SEQUENCE: no

ANTI-SENSE: no

ORIGINAL SOURCE: cDNA placental library

ORGANISM: human

IMMEDIATE EXPERIMENTAL SOURCE: pPlGF-2 (ATCC N.40892)

FEATURES:

from 1 to 321 bp 5' end untranslated region

from 322 to 768 bp coding region

from 769 to 1645 bp 3' end untranslated region

PROPERTIES:

```
         10        20        30        40        50        60
GGGATTCGGGCCGCCCAGCTACGGGAGGACCTGGAGTGGCACTGGGCGCCCGACGGACCA

         70        80        90       100       110       120
TCCCCGGGACCCGCCTGCCCCTCGGCGCCCCGCCCCGCCGGGCCGCTCCCCGTCGGGTTC

        130       140       150       160       170       180
CCCAGCCACAGCCTTACCTACGGGCTCCTGACTCCGCAAGGCTTCCAGAAGATGCTCGAA

        190       200       210       220       230       240
CCACCGGCCGGGGCCTCGGGGCAGCAGTGAGGGAGGCGTCCAGCCCCCCACTCAGCTCTT
```

```
        250       260       270       280       290       300
CTCCTCCTGTGCCAGGGGCTCCCCGGGGGATGAGCATGGTGGTTTTCCCTCGGAGCCCCC

        310       320       330       340       350       360
TGGCTCGGGACGTCTGAGAAGATGCCGGTCATGAGGCTGTTCCCTTGCTTCCTGCAGCTC
                         MetProValMetArgLeuPheProCysPheLeuGlnLeu

        370       380       390       400       410       420
CTGGCCGGGCTGGCGCTGCCTGCTGTGCCCCCCCAGCAGTGGGCCTTGTCTGCTGGGAAC
LeuAlaGlyLeuAlaLeuProAlaValProProGlnGlnTrpAlaLeuSerAlaGlyAsn

        430       440       450       460       470       480
GGCTCGTCAGAGGTGGAAGTGGTACCCTTCCAGGAAGTGTGGGGCCGCAGCTACTGCCGG
GlySerSerGluValGluValValProPheGlnGluValTrpGlyArgSerTyrCysArg

        490       500       510       520       530       540
GCGCTGGAGAGGCTGGTGGACGTCGTGTCCGAGTACCCCAGCGAGGTGGAGCACATGTTC
AlaLeuGluArgLeuValAspValValSerGluTyrProSerGluValGluHisMetPhe

        550       560       570       580       590       600
AGCCCATCCTGTGTCTCCCTGCTGCGCTGCACCGGCTGCTGCGGCGATGAGAATCTGCAC
SerProSerCysValSerLeuLeuArgCysThrGlyCysCysGlyAspGluAsnLeuHis

        610       620       630       640       650       660
TGTGTGCCGGTGGAGACGGCCAATGTCACCATGCAGCTCCTAAAGATCCGTTCTGGGGAC
CysValProValGluThrAlaAsnValThrMetGlnLeuLeuLysIleArgSerGlyAsp

        670       680       690       700       710       720
CGGCCCTCCTACGTGGAGCTGACGTTCTCTCAGCACGTTCGCTGCGAATGCCGGCCTCTG
ArgProSerTyrValGluLeuThrPheSerGlnHisValArgCysGluCysArgProLeu

        730       740       750       760       770       780
CGGGAGAAGATGAAGCCGGAAAGGTGCGGCGATGCTGTTCCCCGGAGGTAACCCACCCCT
ArgGluLysMetLysProGluArgCysGlyAspAlaValProArgArg

        790       800       810       820       830       840
TGGAGGAGAGAGACCCCGCACCCGGCTCGTGTATTTATTACCGTCACACTCTTCAGTGAC

        850       860       870       880       890       900
TCCTGCTGGTACCTGCCCTCTATTTATTAGCCAACTGTTTCCCTGCTGAATGCCTCGCTC

        910       920       930       940       950       960
CCTTCAAGACGAGGGGCAGGGAAGGACAGGACCCTCAGGAATTCAGTGCCTTCAACAACG

        970       980       990      1000      1010      1020
TGAGAGAAAGAGAGAAGCCAGCCACAGACCCCTGGGAGCTTCCGCTTTGAAAGAAGCAAG
```

```
      1030      1040      1050      1060      1070      1080
ACACGTGGCCTCGTGAGGGGCAAGCTAGGCCCCAGAGGCCCTGGAGGTCTCCAGGGGCCT

      1090      1100      1110      1120      1130      1140
GCAGAAGGAAAGAAGGGGGCCCTGCTACCTGTTCTTGGGCCTCAGGCTCTGCACAGACAA

      1150      1160      1170      1180      1190      1200
GCAGCCCTTGCTTTCGGAGCTCCTGTCCAAAGTAGGGATGCGGATTCTGCTGGGGCCGCC

      1210      1220      1230      1240      1250      1260
ACGGCCTGGTGGTGGGAAGGCCGGCAGCGGGCGGAGGGGATTCAGCCACTTCCCCCTCTT

      1270      1280      1290      1300      1310      1320
CTTCTGAAGATCAGAACATTCAGCTCTGGAGAACAGTGGTTGCCTGGGGGCTTTTGCCAC

      1330      1340      1350      1360      1370      1380
TCCTTGTCCCCCGTGATCTCCCCTCACACTTTGCCATTTGCTTGTACTGGGACATTGTTC

      1390      1400      1410      1420      1430      1440
TTTCCGGCCGAGGTGCCACCACCCTGCCCCCACTAAGAGACACATACAGAGTGGGCCCCG

      1450      1460      1470      1480      1490      1500
GGCTGGAGAAAGAGCTGCCTGGATGAGAAACAGCTCAGCCAGTGGGGATGAGGTCACCAG

      1510      1520      1530      1540      1550      1560
GGGAGGAGCCTGTGCGTCCCAGCTGAAGGCAGTGGCAGGGGAGCAGGTTCCCCAAGGGCC

      1570      1580      1590      1600      1610      1620
CTGGCACCCCCACAAGCTGTCCCTGCAGGGCCATCTGACTGCCAAGCCAGATTCTCTTGA

      1630      1640
ATAAAGTATTCTAGTGTGGAAACGC
```

SEQ ID N.2

SEQUENCE TYPE: nucleotide with the corresponding aminoacid sequence for a reading frame

SEQUENCE LENGTH: 828 bp

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: genomic DNA

HYPOTHETICAL SEQUENCE: no

ANTI-SENSE: no

ORIGINAL SOURCE: genomic library

ORGANISM: human

IMMEDIATE EXPERIMENTAL SOURCE: plqfis5

FEATURES:

from 1 to 828 bp 5' end of the fifth intervening sequence of the PlGF gene

from 110 to 172 bp coding region due an alternative splicing of the primary transcript

from 175 to 828 bp coding region due to another alternative splicing of the primary transcript

PROPERTIES:

```
         10        20        30        40        50        60
GTAAGTGGTTTGGCTGGGGCTCGGGGCTATTCTATTCTCGGGCCTGCCAGCCTCTGTCCT

         70        80        90       100       110       120
AGCATGGGGTTCCCCAGCCACCTTGTCCTGACGCTTGGCTTATTGCAGGAGGAGACCCAA
                                                  ArgArgProLy

        130       140       150       160       170       180
GGGCAGGGGGAAGAGGAGGAGAGAGAAGCAGAGACCCACAGACTGCCACCTGTGAGTGCG
sGlyArgGlyLysArgArgArgGluLysGlnArgProThrAspCysHisLeu
```

```
        190       200       210       220       230       240
CGGGGTCCCAGGGATGGCGAGGAGGCTGGGCCCGAGGGGAGCCCCGCCTTGCCGCCAGGG

        250       260       270       280       290       300
TTAGGTTGGGGAGGGGGAGAGGCAGGACTGAGGCCAGTCTTGGGG(G)CAGAACAGGGAN

        310       320       330       340       350       360
CTGCACCTCCTCAAGACTCTAGGGCCCAGGAAGCATCAGTGGACCTTGGTTTTTATCCCG

        370       380       390       400       410       420
GCTTAGCCTAGGTTTCCATTGACCTTCAACAAATCATTTCACCTTTGTCAGCCTAGCTTT

        430       440       450       460       470       480
TCTCTGTGTAGAATGAGGGGCAGGAGGTCCAGCAAACATTCAGTCACTCTACAAACATTT

        490       500       510       520       530       540
ACTGAGCACTTACTGTGTGTCAGGTACATCTGTGAGCAAACAAACAGGATTCCTGCACAT

        550       560       570       580       590       600
TAGTGTTTACCTTTTAGTGATTAAAAGTCTGTCATCAGCTGAGACGTTATCTGGGGCCAC

        610       620       630       640       650       660
TTCCTAGTAGCCCGGGGAACATGTGCCCTCNCACTGTCTCCCAGGAGTATTTTTGCCTGT

        670       680       690       700       710       720
GGGTCCCCTTGCTGCTTCTAACCCACTTCGTACCTTGTGGGCAGCAGAATGGAGCCCCAG

        730       740       750       760       770       780
GCCTGAGTGTGGCTGGGAGAGAAGGATGAGAGGAGGGAAAACCCAAATCTGTGAGAGTAA

        790       800       810       820
ATAGAAAAAATAAAATATTTCACGTGCACAGTCAATCAGTCACTGAAG
```

CLAIMS

1. A nucleotide sequence comprising a sequence coding for a protein with angiogenesis regulative activity having the sequence of SEQ ID N1.

2. A nucleotide sequence according to claim 1 wherein said coded protein derives from alternative splicings, different from those generating the sequence protein of SEQ ID N1.

3. A nucleotide sequence according to claim 2 wherein said alternative splicing generates a transcript coding for a protein with SEQ ID N1, with an insertion of 21 aa., as in SEQ ID N2, at position 141-142 of SEQ ID N1.

4. A nucleotide sequence according to any of the previous claims, wherein said protein lacks one or more amino acids.

5. A nucleotide sequence according to any of the previous claims, wherein said protein has an altered regulative activity of angiogenesis.

6. A nucleotide sequence according to claims 4 or 5 wherein said protein is deleted from the amino acid 1 to the amino acid 31.

7. A nucleotide sequence according to any of claims 1 to 3 wherein one or more amino acids have been substituted in said protein.

8. A nucleotide sequence according to claim 7, wherein said protein shows an altered regulative activity of angiogenesis.

9. A nucleotide sequence according to any of the previous claims wherein said protein is an allelic

derivative.

10. A nucleotide sequence complementary to a sequence according to any of the previous claims.

11. A nucleotide sequence according to any of the previous claims, to which a nucleotide sequence, that can be translated with the same reading frame, has been covalently linked to the 5' or to the 3' positions.

12. A nucleotide sequence according to claim 11, wherein said linked nucleotide sequence codes for an amino acid sequence which does not interfere with the regulative activity of angiogenesis of the protein.

13. A nucleotide sequence according to claims 11 or 12 wherein said linked nucleotide sequence codes for a protein having a toxic activity.

14. A nucleotide sequence comprising the sequence of SEQ ID N1, wherein the 5'end region from the nucleotide 1 to the nucleotide 321 is untranslated, the region from the nucleotide 322 to the nucleotide 768 codes for a protein with angiogenesis regulative activity, the 3' end region is untranslated.

15. A nucleotide sequence according to claim 14 comprising at least a part of sequence of SEQ ID N2.

16. A nucleotide sequence according to claim 14 or 15 lacking one or more nucleotides, coding for a protein having a regulative activity of angiogenesis.

17. A nucleotide sequence according to claims 14 or 15 lacking one or more nucleotides, coding for a protein having an altered regulative activity of angiogenesis.

18. A nucleotide sequence according to claims 16 or 17 which is deleted from the nucleotide 322 to the

nucleotide 414.

19. A nucleotide sequence according to claims 14 or 15 wherein one or more nucleotides have been substituted, coding for a protein having a regulative activity of angiogenesis.

20. A nucleotide sequence according to claims 14 or 15 wherein one or more nucleotides have been substituted, coding for a protein with an altered regulative activity of angiogenesis.

21. A nucleotide sequence according to claims 14 or 15, wherein said untranslated 5' region comprises a regulative region of the translation.

22. A nucleotide sequence according to claim 21, wherein said regulative region forms a stem-loop secondary structure.

23. A nucleotide sequence which is allelic to the nucleotide sequence according to claims 14 or 15.

24. A nucleotide sequence which is complementary to a nucleotide sequence according to any of the previous claims from 14 to 23.

25. A protein comprising the amino acid sequence of SEQ ID N1 having a regulative activity of angiogenesis.

26. A protein according to claim 25 wherein said protein comprises an insertion derived from alternative splicing of the primary transcript.

27. A protein according to claim 26 wherein said insertion is at position 141-142 of SEQ ID N1 and comprises a sequence coded by at least a portion of SEQ ID N2.

28. A protein according to claim 27 wherein said

sequence is of 21 amino acids as in SEQ ID N2.

29. A protein according to any of claims 22 from 25 to 28 wherein said protein lacks one or more amino acids.

30. A protein according to claim 29, wherein said protein shows an altered regulative activity of angiogenesis.

31. A protein according to claims 29 or 30, wherein said protein is deleted from the amino acid 1 to the amino acid 31.

32. A protein according to any of claims from 25 to 28 wherein one or more amino acids have been substituted in said protein.

33. A protein according to claim 32 wherein said protein shows an altered regulative activity of angiogenesis.

34. A protein coded by a nucleotide sequence which is an allelic derivative of the nucleotide sequence according to any of claims from 1 to 3.

35. A protein comprising an amino acid sequence according to anyone of the previous claims from 25 to 34, to which an amino acid chain is covalently linked at the terminal COOH or NH2 groups.

36. A protein according to claim 35, wherein said amino acid chain does not interfere with the regulative activity of the angiogenesis of the protein itself.

37. A protein according to any of claims 35 or 36, wherein said amino acid chain shows a toxic cell activity.

38. A vector comprising: a) the replication origin of a bacterial plasmid, b) a selective marker, c) a

promoter and under the control of said promoter, d) the nucleotide sequence according to any of the previous claims from 1 to 24.

39. A vector according to claim 38, wherein said selective marker is a gene coding for an antibiotic resistance.

40. A vector according to any of claims 38 or 39, wherein said promoter is T7 phage RNA polymerase promoter.

41. A vector according to any of claims from 38 to 40 wherein said vector comprises the plasmid pGem 1.

42. A vector according to any of claims from 38 to 40 wherein said vector comprises the plasmid pET3.

43. A vector according to any of claims from 38 to 42 comprising one or more sequences that regulate the transcription.

44. A vector according to claim 43, wherein said sequences comprise a transcription enhancer.

45. A vector according to claim 43 wherein said sequences comprise an inducible promoter.

46. A vector according to any of claims from 38 to 45 comprising a polyadenylation site.

47. Cells transformed with a vector according to any of claims from 38 to 46.

48. Cells according to claim 47 wherein said cells are bacterial cells.

49. Cells according to claim 48 wherein said bacteria are E.coli.

50. Cells according to claim 47 wherein said cells are eukaryotic cells.

51. A process for the production and extraction of the protein PlGF from cells according to any of claims from 47 to 49, said process comprising: growing the bacterial culture in liquid medium to an optical density between 0.2 and resuspending in TE; centrifuging and resuspending in a lysis buffer; lysing the cells.

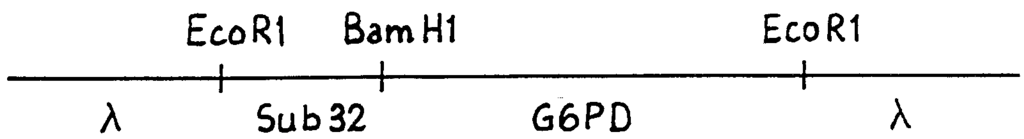

EcoR1
BamH1
EcoR1
λ
Sub32
G6PD
λ

FIG. 1

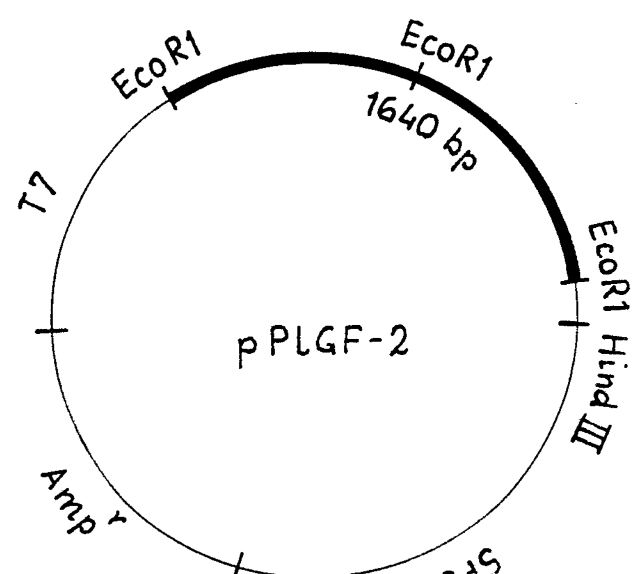

EcoR1
EcoR1
1640 bp
EcoR1
Hind III
T7
pPLGF-2
Amp r
Sp6

FIG. 3

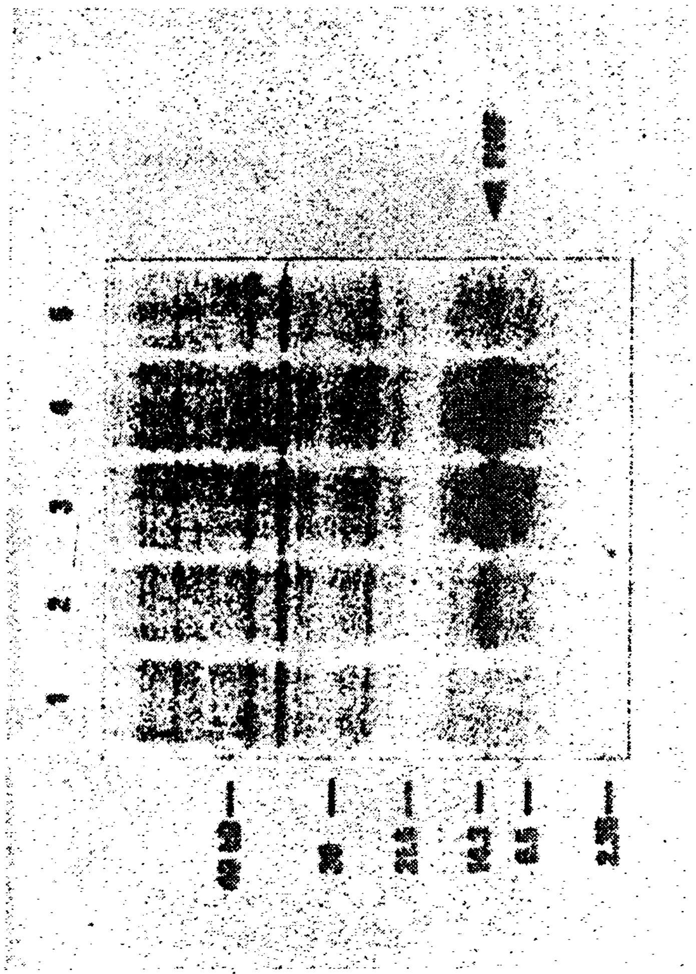


FIG. 5

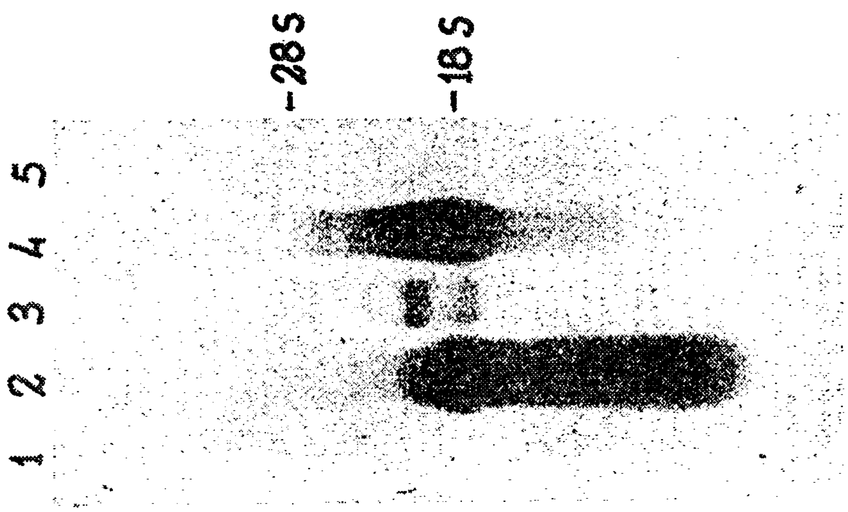


FIG. 2

CONSTRUCTION OF pET-PLGF

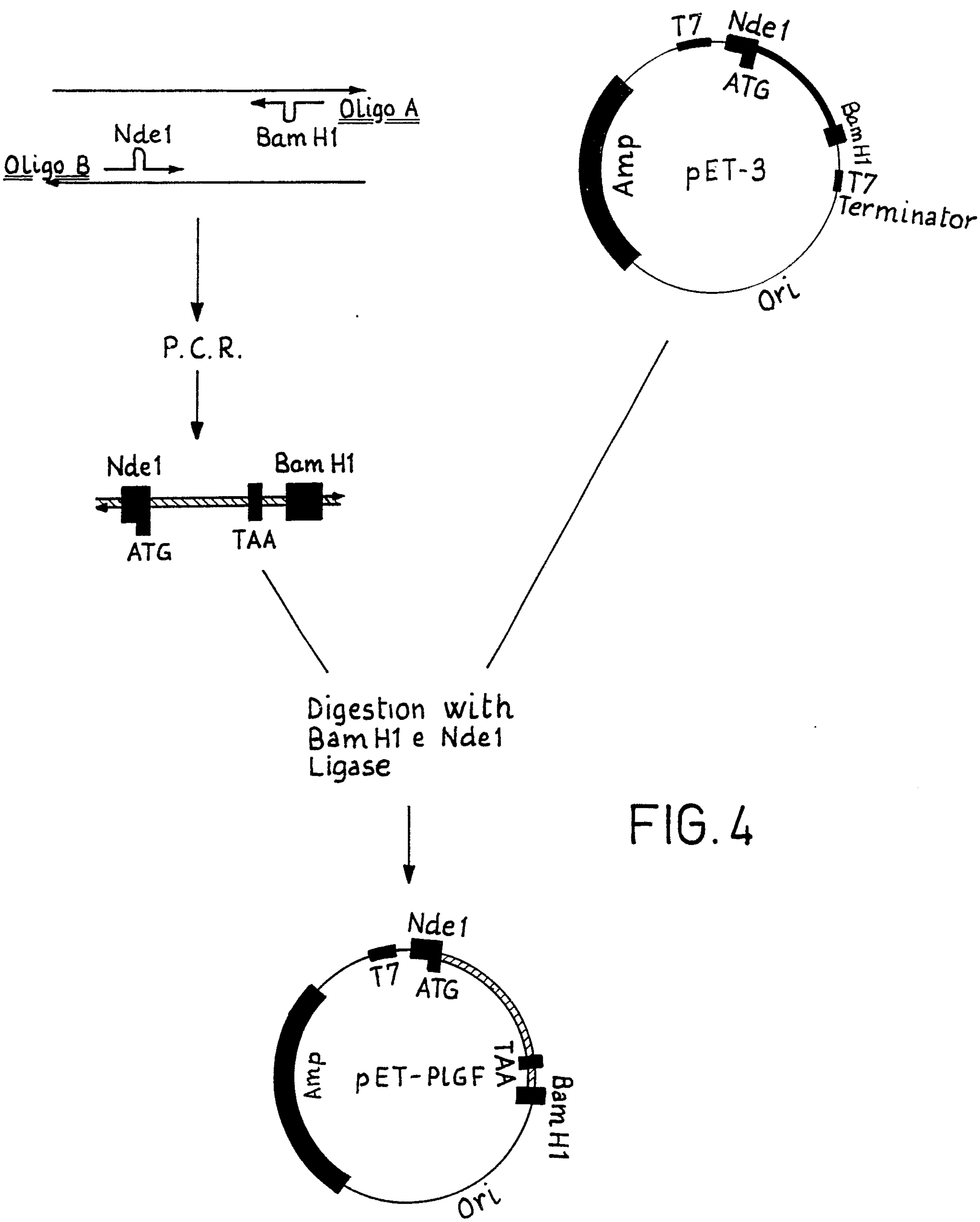


FIG. 4

International Application No PCT/IT 91/00079

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)[6]

According to International Patent Classification (IPC) or to both National Classification and IPC

Int.Cl. 5 C12N15/16; C12N15/62; C12P21/02; C12N1/21
C12N5/16; //(C12N1/21;C12R1:19)

II. FIELDS SEARCHED

Minimum Documentation Searched[7]

Classification System	Classification Symbols
Int.Cl. 5	C07K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched[8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT[9]

Category °	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No.[13]
P,X, O	J. CELL. BIOCHEM. SUPPL. 0 (15 PART F) 1991, NEW YORK, US; ABSTRACT NR. CF 316 page 242; M. G. PERSICO ET AL.: 'A NOVEL HUMAN ANGIOGENIC FACTOR PLGF' see the whole document & 'MEETING ON FGF, ENDOTHELIAL CELL GROWTH FACTORS AND ANGIOGENESIS HELD AT THE 20TH ANNUAL MEETING OF THE KEYSTONE SYMPOSIA ON MOLECULAR AND CELLULAR BIOLOGY, KEYSTONE, COLORADO, USA, APRIL 1-7 1991'	1,10, 14-16, 21-22, 24-25

° Special categories of cited documents : [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
06 DECEMBER 1991	16.12.91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	THIELE U.H.-C.H.

3